

Understanding the Future of Video Coding

Yucel Altunbasak
Georgia Institute
of Technology

Hasan Ates
Sabanci University

Iain E.G. Richardson, *H.264 and MPEG-4 Video Compression: Video Coding for Next-Generation Multimedia*, John Wiley & Sons, 2003, \$102.00, 206 pp., ISBN 0-470-84837-5.

In today's world, a wide range of multimedia services has become an integral part of many industries—from telecommunications to broadcasting and entertainment to consumer electronics. Through ongoing research and technological developments, the scope and the variety of multimedia applications and products are widening quickly. This constant change brings its own set of problems: mainly, the difficulty of achieving compatibility and interoperability among different systems. When different consumer products share multimedia data, this information has to be represented and coded using compatible syntax, semantics, and tools. The standardization process aims to solve this problem, and video coding standards are at the forefront of this effort.

Overview

Iain Richardson's book focuses on two leading candidates for future video coding standards: MPEG-4 Visual (MPEG-4 Part 2) and H.264 (also known as MPEG-4 Part 10). Standardization is a difficult and time-consuming task. When the standard is finalized and published, the specifications are hardly accessible even for a technical audience. Reference books such as this one are written to describe the various components of the standard in more detail and to look at the design issues from the encoder's perspective, which isn't included in the standard.

In Richardson's book, the author introduces the subject by following a clear, practical, and informative explanation of the two standards (MPEG-4 Visual and H.264). The reader gets a simplified understanding of core concepts and how to implement them. The author keeps his

focus throughout the book and deals with complex issues only to the extent that it helps the reader understand the standards.

Video coding is indeed a complex and extensive subject. Because of the widespread demand for technologies—such as digital high-definition TV, streaming video over the Internet, and wireless video—new emerging standards try to provide a number of tools and features that could support a variety of applications and specifications. Hence, both MPEG-4 Visual and H.264 contain a rich set of coding tools organized into profiles, specifically designed for different types of applications. This book follows a rather practical strategy and explains the details of each profile, going from the basic to the most advanced. The basic coding tools are described in greater detail, and additional features are provided as supplemental information. This application-oriented approach makes the book easier to read and refer to.

Getting started

The book starts with a short introduction to digital video coding. Chapter 2 deals with digital video formats, color spaces, and quality measures. Chapter 3 introduces a hybrid video codec as the generic model for all major video coding standards. This model is divided into its basic building blocks, namely block-based motion compensation, transform coding, quantization, and entropy coding. The author explains these topics with minimal detail. The logical and historical development of video coding is mostly left out, including only the essentials necessary for the future chapters.

Chapter 4 gives an overview of the MPEG-4 Visual and H.264 standards, and describes the role of expert groups, ISO MPEG and ITU VCEG (Motion Picture Experts Group and Video Coding Experts Group), and their working schedule during the standardization process. For a novice, it

continued on p. 87

continued from p. 88

may be interesting to read how the features of the standards are modified and finalized as a result of collective efforts of many groups from research institutes and companies. This chapter also summarizes the contents of the MPEG-4 Visual and H.264 standard documents and highlights some important differences between the two. Previous video coding standards, such as H.263 and MPEG-2, are only briefly mentioned. The lack of further information on these topics makes it hard for an audience new to the field to grasp the theoretical and practical reasoning behind different implementation choices for each standard.

Advancing the options

Chapters 5 and 6 are dedicated to MPEG-4 Visual and H.264, respectively. The book focuses on the simple and core profiles for MPEG-4 and on the baseline profile for H.264. These profiles cover the most basic coding tools of each standard. For other profiles, only the additional features are described.

In comparing the two standards, MPEG-4 Visual is singled out for its flexibility in coding various types of visual data, such as arbitrarily shaped regions of a scene, and H.264 is acclaimed for its coding efficiency and reliability. The author explains in clear and simple terms the essential features responsible for the widespread acknowledgment of each standard as the future of video coding. Difficult concepts—such as shape coding, scalable coding in MPEG-4, or B slices and weighted prediction in H.264—are handled with ease, neither being too technical nor too sketchy.

As in other video coding books, these chapters approach the coding process from the encoder's perspective. This helps not only to understand the standards better, but also to get a sense of how to design a good encoder.

In Chapter 5, the author covers MPEG-4 Visual under two main parts: coding rectangular frames and arbitrary shaped regions. Scalable coding, texture coding, coding studio-quality video, and coding synthetic visual scenes are among the topics that are only briefly mentioned. It's especially stimulating to read about the innovative toolkits used for object-based coding and to grasp the level of coding flexibility achieved by MPEG-4 Visual.

In Chapter 6, Richardson discusses H.264's implementation details with more technical depth than MPEG-4 Visual. He carefully follows the implementation of each of the building

blocks, going through different inter- and intra-prediction modes, the deblocking filter, transform and quantization, and entropy coding. Almost all of the technical details are provided without confusing or boring the reader. For instance, the author explains B-slice coding and reference picture selection with a level of clarity that's hard to get from the standard document.

Richardson further describes the practical issues related to the design of building blocks of a video encoder in Chapter 7. These details provide the nontechnical audience with a sense of how the standards are implemented in software and hardware. This chapter also compares the performance of the two standards—based on subjective and objective quality measures and computational costs—and emphasizes once again the superior coding efficiency of H.264. Some of the optional features are evaluated and compared in terms of their contribution to the coding performance. The author also provides a valuable discussion about how rate-control algorithms can reduce bit rate variations in the coded sequence, and hence decrease input and output buffer sizes and coding delay.

Further considerations

Chapter 8 examines the requirements of some current and emerging applications and explains how these two standards might emerge as the market winner, depending on the various driving forces in the industry. The author emphasizes that performance is only one of the factors, and industry support, costs, and availability of development tools are just as important. The majority of commercial MPEG-4 Visual codecs target broadcast-quality or streaming video applications. Initial target applications for H.264 are singled out to be broadcast TV (high-definition TV included) and DVD, with mobile applications probably emerging afterward.

In summary, the book is a well-written and practical reference guide to these standards, and the author's fluid style turns what might be a complex subject into enjoyable reading that's even accessible to a nontechnical audience. However, since difficult concepts are explained with such clear and concise language but without an in-depth analysis, the book leaves the reader wanting more information on the topic. **MM**

Readers may contact Yucel Altunbasak and Hasan Ates at yucel@ece.gatech.edu and hasanates@sabanciuniv.edu.